THE EFFECT OF INVERSE SAMPLING ON RANKING

MULTINOMIAL CELL PROBABILITIES*

by

Theophilos Cacoullos and Milton Sobel

Technical Report No. 55

UNIVERSITY OF MINNESOTA

DEPARTMENT OF STATISTICS

April, 1965

THE EFFECT OF INVERSE SAMPLING ON RANKING MULTINOMIAL CELL PROBABILITIES*

by

Theophilos Cacoullos and Milton Sobel

Technical Report No. 55

University of Minnesota

Minneapolis, Minnesota

*This report was supported by the National Science Foundation under Grant GP-3813 and by the United States Air Force Office of Scientific Research - Office of Aerospace Research.

Errata Sheet for

The Effect of Inverse Sampling on Ranking Multinomial Cell Probabilities

Page 4 line 13. Blurred part of equation ([illegible]) should [illegible]

Page 6 line 8. Blurred part should read "p [illegible] /(1 [illegible])"

Page 6 line 14. Blurred part should read "[illegible]".

Page 6 line 16. Last part should read "[illegible]".

Table I, $P^* = .90$, $m = 1$, [illegible] 1.2. Change entry "99" to [illegible]

Table II, $P^* = .9$[illegible], $m = 1$, [illegible] Change six entries to read "1[illegible], 181.99, ([illegible].04), 188.73, ([illegible]) and 199", [illegible]

Table II, $P^* = .75$, $m = 1$, [illegible] 1.[illegible] Change "[illegible]" to "[illegible]".

Table II, $P^* = .75$, $m = 3$, [illegible] 1.6. Change entry for a_n from "[illegible]" to "[illegible]"

Table II. Change first sentence of footnote to read

"All entries in parentheses are asymptotic; all entries without parentheses are based on exact formulas (and hence [illegible] the integer [illegible] in the same cell is correct), except [illegible] for $m = 2, 3$ is based on the approximation ([illegible]9)."

Table II, $P^* = .90$, $m = 3$, [illegible] Put "[illegible]" on entry "[illegible]" and add a second footnote to read "This entry is based on the [illegible] formula ([illegible]) the result obtained by using the asymptotic formula [illegible] ([illegible]5)".

Dedicated

to

STEFAN BANACH

whose "matches", even behind barbed wires,

lighted the torch for new areas of research.

THE EFFECT OF INVERSE SAMPLING ON RANKING
MULTINOMIAL CELL PROBABILITIES

by

Theophilos Cacoullos and Milton Sobel

1. Summary

The problem of selecting the particular one of k multinomial cells with the highest probability is considered from the ranking theory--indifference zone point of view. The sampling procedure used differs from the fixed sample size procedure (FSP) of Bechhofer, Elmaghraby and Morse [1] in that observations are taken one at a time until any one cell has N counts in it. It is shown that for $k = 2$ the same requirement on the probability of a correct selection can be satisfied by the inverse sampling procedure (ISP) with a smaller expected total number of observations than the fixed sample size of [1], regardless of the true parameter point. A concept of asymptotic proportion saved (APS_2) by the ISP relative to the FSP is defined as an asymptotic limit as the probability requirement gets stricter, i.e., $\delta^* \to 1$ (cf. (3.2) and (7.6)), and it is shown to be positive for any $P^* > \frac{1}{2}$ and any k; analogous comparisons are made for the worst case when all the cells have equal probability.

The required value of N for this problem is easily obtained from existing tables [4] and relevant asymptotic results obtained here; a short table of such values is included.

2. Introduction

This paper was motivated by the Banach match box problem [3]. In this problem matches are drawn from either of two boxes (of N matches each) with specified probabilities until one box is found to be empty; the problem deals with the distribution of the number R remaining in the other box; here R takes on integer values from 0 to N, inclusive. A closely related problem is to stop sampling when the last match is drawn out from either box, so that R only takes on values from 1 to N, inclusive. This paper deals mostly with generalizations and applications of the latter variation, which appears to have more application than the original variation of the problem.

Although the problem deals with the removal of matches from the boxes, we could equally well consider the same problem with matches being put into the boxes. Then we continue until either box has N matches in it and consider the distribution of the number X of matches in the other box. Clearly the two problems are equivalent with X corresponding to N - R. This latter aspect shows that the problem is one of inverse sampling from a Bernoulli or, more generally for $k \geqq 2$, from a multinomial distribution.

The problem of deciding which cell has the highest probability is considered here with this inverse sampling procedure (ISP). The corresponding problem with a fixed sample size procedure (FSP) was considered by Bechhofer, Elmaghraby and Morse [1]. The two procedures are comparable since they satisfy the same requirement on the probability of a correct selection and the so-called least favorable configuration in the parameter space is the same for both procedures.

Let n_o denote the fixed sample size needed for the FSP and let $E(T|LF)$ denote the expected total number of observations required from the ISP in the least favorable configuration (defined in Section 4). The main results of this paper deal with a comparison of n_o and $E(T|LF)$; for $k = 2$ we can make exact comparisons and for $k \geqq 3$ we make comparisons based on asymptotic normal approximations. In [1] the normal approximation to n_o is carried out after the arc sine-square root transformation is applied to the chance variables; a corresponding (but not the same) transformation is used in this paper to give a better normal approximation. Comparisons of n_o and $E\{T|W\}$, i.e., the expected value of T in the equal parameter (W) configuration, are also made in this paper.

The value of N needed to make the ISP explicit and to satisfy the probability requirement turns out to be easily obtainable from an existing table in [4]; a short table of N- values is included here to make this paper self-contained. A basic tool used several times in this paper is an identity (see (4.2) below) connecting a negative multinomial sum and a Dirichlet integral; this identity is proved in [7].

3. Formulation of the Problem

Observations are taken one at a time from a multinomial distribution with k cells, $C_1, C_2, \ldots, C_k$ with probabilities $p_1, p_2, \ldots, p_k$, respectively, such that $\sum_{1}^{k} p_i = 1$. The ordered values of these cell probabilities are denoted by

$$p_{[1]} \leq p_{[2]} \leq \cdots \leq p_{[k-1]} \leq p_{[k]} \quad , \tag{3.1}$$

and we let $C_{(i)}$ denote the cell associated with $p_{[i]}$ $(i = 1, 2, \ldots, k)$. The problem (or goal) is to select the cell $C_{(k)}$, which we shall sometimes refer to as the "best" cell. The probability requirement to be satisfied is expressed in terms of the ratio $\delta_{k,k-1} = \delta(\text{say})$ of $p_{[k]}$ to $p_{[k-1]}$ and two preassigned constants, P^* and δ^*, such that $1/k < P^* < 1$ and $\delta^* > 1$. For any $\delta > 1$, we define a correct selection (CS) in the obvious way, as the selection of the cell $C_{(k)}$; for $\delta = 1$ we would define it as the selection of any cell $C_{(i)}$ with $p_{[i]} = p_{[k]}$, but we shall not need the definition for $\delta = 1$. The experimenter would like to have a procedure R for selecting the best cell which satisfies the probability requirement,

$$P\{CS|R\} \geq P^* \qquad \text{for } \delta \geq \delta^* \quad . \tag{3.2}$$

Let $R_o = R_o(k, \delta^*, P^*)$ denote the fixed sample procedure (FSP) of [1] that satisfies (3.2) and let $R_1 = R_1(k, \delta^*, P^*)$ denote the following inverse sampling procedure (ISP) that satisfies (3.2). Let $E_i = E_i^{(N)}$ denote the event that we observe N observations from cell C_i before observing N observations from any other cell ($i = 1, 2, \ldots, k$); the integer $N = N(k, P^*, \delta^*)$ is chosen in advance of experimentation in such a way that (3.2) is satisfied.

Inverse Sampling Procedure (R_1): Continue sampling one at a time until one of the events E_i occurs and then select the cell C_i as being the best cell.

4. The $P\{CS|R_1\}$ and the Determination of N

Let $E_{(j)} = E_{(j)}^{(N)}$ denote the event that the cell $C_{(j)}$ associated with $p_{[j]}$ is emptied first ($j = 1, 2, \ldots, k$) and let $X_{(i)}$ denote the number of observations taken from $C_{(i)}$, $i \neq j$ at the time $E_{(j)}$ occurs, so that $0 \leq X_{(i)} \leq N-1$ for each $i \neq j$. [In the dual model of taking out items (without replacement) from k cells all containing N items at the outset, we use $Z_{(i)}$ to denote the number remaining in cell $C_{(i)}$ for $i \neq j$ when $E_{(j)}$ occurs; then $Z_{(i)} = N - X_{(i)}$ and $1 \leq Z_{(i)} \leq N$.]

The probability of a correct selection is easily seen to be

$$(4.1) \qquad P\{CS|R_1\} = P\{E_{(k)}\} = p_{[k]}^N \sum_{x_1=0}^{N-1} \cdots \sum_{x_m=0}^{N-1} \frac{\Gamma(N+x_o)}{\Gamma(N)} \prod_{i=1}^{m} \frac{p_{[i]}^{x_i}}{x_i!} ,$$

where $m = k-1$ and $x_o = x_1 + x_2 + \ldots + x_m$. The multiple sum in (4.1) can be written as a Dirichlet integral using theorem 2.4 of [7] : if $s_1, s_2, \ldots, s_m$ are positive integers, $s > 0$, $\theta_i \geq 0$, $i = 1, 2, \ldots, m$ and $\theta_o = \theta_1+\theta_2+\ldots+\theta_m < 1$, then

$$(4.2) \qquad (1-\theta_o)^s \sum_{x_1=0}^{s_1-1} \cdots \sum_{x_m=0}^{s_m-1} \frac{\Gamma(s+x_o)}{\Gamma(s)} \prod_{i=1}^{m} \frac{\theta_i^{x_i}}{x_i!}$$

$$= \frac{\Gamma(s+s_o)}{\Gamma(s)\prod_{i=1}^{m}\Gamma(s_i)} \int_{\theta_1^*}^{\infty} \cdots \int_{\theta_m^*}^{\infty} \frac{\prod_{i=1}^{m}\{y_i^{s_i-1}dy_i\}}{(1+y_o)^{s+s_o}} ,$$

where $y_o = y_1+y_2+\ldots+y_m$, $s_o = s_1+s_2+\ldots+s_m$, $\theta_i^* = \theta_i/(1-\theta_o)$, $i = 1, 2, \ldots, m$ and x_o is as defined above. Hence we obtain from (4.1) with $m = k-1$

$$(4.3) \qquad P\{CS|R_1\} = \frac{\Gamma(kN)}{\{\Gamma(N)\}^k} \int_{\delta_{k,1}^{-1}}^{\infty} \cdots \int_{\delta_{k,m}^{-1}}^{\infty} \frac{\prod_{i=1}^{m}\{y_i^{N-1}dy_i\}}{(1+y_o)^{kN}} ,$$

where $\delta_{k,i} = p_{[k]}/p_{[i]}$ $(i = 1, 2, \ldots, m)$.

[It should be noted that the probability $P\{C_{(i)}|R_1\}$ that the cell $C_{(i)}$, associated with $p_{[i]}$, empties first is given by the same integral as in (4.3) except that the lower limits are replaced by $\delta_{i,j}^{-1}$ $(j = 1,2,\ldots,i-1,i+1,\ldots,k)$.]

Clearly the $P\{CS|R_1\}$ is minimized if the lower limits of integration in (4.3) are maximized or the $\delta_{k,i}$ are minimized subject to the condition $\delta_{k,i} \geq \delta^*$.(cf.(3.2)). Hence we need only set $\delta_{k,i} = \delta^*(i=1,2,\ldots,m)$, i.e., the least favorable configuration (LFC), where the $P\{CS|R_1\}$ attains its minimum subject to the condition $\delta \geq \delta^*$, is given by

$$(4.4)\qquad \begin{aligned} p_{[1]} &= p_{[2]} = \cdots = p_{[m]} = \frac{1}{\delta^*+m} = q(\text{say}) \\ p_{[k]} &= \frac{\delta^*}{\delta^*+m} = p(\text{say})\ ;\quad p + m\,q = 1 \quad . \end{aligned}$$

We remark that the same configuration (4.4) is least favorable for the FSP in [1], the proof in that case being given in a separate paper [5]. This makes the FSP and the ISP directly comparable and comparisons are made in section 7 below. The above discussion proves

<u>Theorem 4.1</u> The value of N required to satisfy (3.2) is the smallest integer equal to or greater than the solution $n = n(k,\delta^*,P^*)$ of the equation

$$(4.5)\qquad \frac{\Gamma(kn)}{[\Gamma(n)]^k} \int_{1/\delta^*}^{\infty} \cdots \int_{1/\delta^*}^{\infty} \frac{\prod_{i=1}^{m} [y_i^{n-1} dy_i]}{(1+y_o)^{kn}} = P^* \quad .$$

Using already existing tables of $1/\delta^*$- values satisfying (4.5) for selected values of k, P* and n given in [4], it is easy to find the required N- values for selected values of k, P* and δ^*; our symbols $k = m+1$, N P* and δ* correspond to $k = p+1$, $\nu/2$, P* and $1/c$, respectively, in [4]. (See Table I below.)

Since these tables are limited to $N \leq 25$ and apply only for P*=.75, .90, .95, and .99 it is desirable for both practical and theoretical reasons to develop an asymptotic expression for the left side of (4.5) and an asymptotic solution of

(4.5) for n. We shall also need asymptotic expressions for $E(T|LF)$ and $E(T|W)$ for the ISP to make comparisons with the value of n_o needed by the FSP since the exact expressions for $k > 2$ are involved. Finally, these asymptotic results may have some interest per se.

For the special case k = 2 the right side of (4.3) can be written as an Incomplete Beta function by applying the transformation $1+y_1 = u^{-1}$; we obtain

$$P\{CS|R_1\} = \frac{\Gamma(2N)}{[\Gamma(N)]^2} \int_0^p u^{N-1}(1-u)^{N-1}du = I_p(N,N) \tag{4.6}$$

where $p = \delta^* /(1+\delta^*)$. This result will be expecially useful in making comparisons in Section 7 with the results for the FSP in [1].

5. Asymptotic Theory

Consider the random vector $Y = (Y_1, Y_2, \ldots, Y_m)$ with the Dirichlet density

$$f(y_1,\ldots,y_m) = \frac{\Gamma(s+s_o)}{\Gamma(s)\prod_{i=1}^{m}\Gamma(s_i)} \frac{\prod_{i=1}^{m} y_i^{s_i-1}}{(1+y_o)^{s+s_o}} \qquad \begin{matrix} y_i \geq 0 \\ (i = 1,2,\ldots,m) \end{matrix} \tag{5.1}$$

where s and the s_i are any positive numbers and $s_o = s_1+s_2+\ldots+s_m$. The means of the Y_i for $s > 1$ and the covariances of the Y_i for $s > 2$ are easily computed and if s and the $s_i (i = 1,2,\ldots,m) \to \infty$ in such a way that $s/N \to 1$ and

$$\frac{s_i}{s} \to \lambda_i \qquad (i = 1,2,\ldots,m,) \tag{5.2}$$

where the λ_i are positive, finite limits, then

$$E(Y_i) = \frac{s_i}{s-1} \doteq \lambda_i \qquad (i = 1,2,\ldots,m) \quad ,$$

$$\mathrm{Var}(Y_i) = \frac{s_i(s_i+s-1)}{(s-1)^2(s-2)} \doteq \frac{\lambda_i(1+\lambda_i)}{N} \qquad (i = 1,2,\ldots,m) \quad , \tag{5.3}$$

$$\rho(Y_i,Y_j) = \sqrt{\frac{s_i s_j}{(s_i+s-1)(s_j+s-1)}} \doteq \sqrt{\frac{\lambda_i\lambda_j}{(1+\lambda_i)(1+\lambda_j)}} \qquad (j \neq i) \quad .$$

It is clear from the left side of (4.5) that our main interest is in the case $\lambda_i = 1$ $(i = 1,2,\ldots,m)$, in which case the common mean, variance and correlation in (5.3) are asymptotically 1, 2/N and 1/2, respectively.

Theorem 5.1: If the limits λ_i in (5.2) are positive and finite then the asymptotic distribution of the variables

$$(5.4) \qquad X_i = \sqrt{\frac{N}{\lambda_i(1+\lambda_i)}} \quad (Y_i - \lambda_i) \qquad (i = 1,2,\ldots,m)$$

is a joint normal distribution with zero means, unit variances and an $m \times m$ correlation matrix $\Lambda = \{\rho_{ij}\}$ with $\rho_{ij} = \rho(Y_i, Y_j)$ given by (5.3). In particular, for $\lambda_i = 1$ $(i = 1,2,\ldots,m)$ we obtain

$$(5.5) \quad f(x_1, \ldots, x_m) = \frac{\exp\left\{-\left[m\sum_1^m x_i^2 - 2\sum_{i<j} x_i x_j\right]/(m+1)\right\}}{\pi^{m/2}\sqrt{m+1}}$$

Proof: Using Stirling's well-known approximation to $\Gamma(\cdot)$ throughout (5.1) and some elementary limiting arguments, we obtain theorem 5.1 by a direct limiting process; the details are omitted.

Corollary 5.1: If $\lambda_i = \lambda > 0$ $(i = 1,2,\ldots,m)$ and $\lambda\,\delta^* \neq 1$, then

$$(5.6) \qquad N \sim \frac{(\delta^*)^2\lambda(1+\lambda)}{(\lambda\delta^*-1)^2} \, H_m^2\left(P^*, \frac{\lambda}{1+\lambda}\right)$$

where $H = H_m(P^*,\rho)$ is given in terms of the standard univariate normal density $f(x)$ and the corresponding c.d.f. $F(x)$, as the solution of

$$(5.7) \qquad \int_{-\infty}^{\infty} F^m\left(\frac{x\sqrt{\rho}+H}{\sqrt{1-\rho}}\right) f(x)\,dx = P^* \quad .$$

Proof: If N is large then, by theorem 5.1, N satisfies

$$(5.8) \qquad P\left\{X_i \geq -\frac{\sqrt{N}\,(\lambda\delta^*-1)}{\delta^*\sqrt{\lambda(1+\lambda)}} \quad (i = 1,2,\ldots,m)\right\} = P^* \quad ,$$

where the X_i are standard normal chance variables with common correlation $\rho = \lambda/(1+\lambda) > 0$. It is clear that for $\rho > 0$ we can set

(5.9) $$X_i = \sqrt{1-\rho}\, Y_i + \sqrt{\rho}\; Y_o \qquad (i = 1,2,\ldots,m)$$

where the Y_i and Y_o are all independent, standard normal chance variables. Hence (5.8) can be written in the form (5.7) with $\rho = \lambda/(1+\lambda)$ and

(5.10) $$H = H_m(P^*, \frac{\lambda}{1+\lambda}) = \frac{\sqrt{N}(\lambda\delta^*-1)}{\delta^*\sqrt{\lambda(1+\lambda)}} .$$

Solving (5.10) for N gives the result (5.6), which completes the proof.

It was found that a better normal approximation could be had if we first transform the Dirichlet variable by a logarithmic transformation; this would be the appropriate transformation for stabilizing the variance if a common scale parameter were present. (All logs in this paper are taken to the natural base e.)

Corollary 5.2: If the limits λ_i in (5.2) are positive and finite, then the asymptotic ($N \to \infty$) distribution of the chance variables

(5.11) $$W_i = \sqrt{\frac{N\lambda_i}{1+\lambda_i}}\ \log\,(Y_i/\lambda_i) \qquad (i = 1,2,\ldots,m)$$

is a joint normal distribution with zero means, unit variances and the same $m \times m$ correlation matrix Λ as given by (5.3).

Proof: The proof follows from straightforward asymptotic methods using theorem 5.1 and the facts that asymptotically

$$E(\log Y_i) \sim \log E(Y_i) \sim \log \lambda_i \qquad (i = 1,2,\ldots,m)$$

(5.12) $$\operatorname{Var}\,(\log Y_i) \sim \frac{1}{\lambda_i^2}\ \operatorname{Var}\,(Y_i) \sim \frac{1+\lambda_i}{N\lambda_i} \qquad (i = 1,2,\ldots,m) ,$$

$$\rho(\log Y_i, \log Y_j) \sim \rho(Y_i,Y_j) \sim \sqrt{\frac{\lambda_i\lambda_j}{(1+\lambda_i)(1+\lambda_j)}} \qquad (j \neq i) .$$

From corollary 5.2 we now obtain as in (5.6)

Corollary 5.3: If $\lambda_i = \lambda > 0$ $(i = 1,2,\ldots,m)$ and $\lambda\,\delta^* \neq 1$, then

(5.13) $$N \sim \frac{(1+\lambda)}{\lambda(\log \lambda\delta^*)^2}\ H_m^2\,(P^*, \frac{\lambda}{1+\lambda}) ,$$

where $H = H_m(P^*,\rho)$ is given by (5.7). Tables of H-values have been obtained by several different authors and are available in the literature. For the case $\lambda=1$ of special interest here, it has been found by empirical comparisons with the exact solutions in [4] that the addition of log m, i.e.,

$$(5.14) \qquad N \overset{\cdot}{\approx} \frac{2}{(\log \delta^*)^2} H_m^2 (P^*, \tfrac{1}{2}) + \log m,$$

appears to give a uniform improvement to (5.13) with $\lambda = 1$; some numerical results based on this correction term are given above the heavy lines in Table I.

The following asymptotic results describe the basic structure of the inverse sampling procedure for large values of N. They deal with the asymptotic $(N\to\infty)$ distributions of the number of observations $X_{(i)}$ taken from the cell $C_{(i)}$ associated with the cell probability $p_{[i]}$ when event $E_{(j)}$ occurs $(j \neq i)$ [or equivalently with the number of remaining items $Z_{(i)} = N-X_{(i)}$ in the dual model when $E_{(j)}$ occurs $(j \neq i)$]. We obtain two different types of results, both of interest, according to whether $p_{[k-1]} < p_{[k]}$ or $p_{[1]} = \dots = p_{[k]} = 1/k$; the other cases are not considered. For convenience, let $\theta_i = p_{[i]}/p_{[k]}$ $(i = 1,2,\dots,m)$; we define standardized variables X_i^* and analogous nonnegative variables Z_i^* by

$$(5.15) \qquad X_i^* = (X_{(i)} - N\theta_i)/\sqrt{N\theta_i(1+\theta_i)} \qquad (i = 1,2,\dots,m) \quad ,$$

$$(5.16) \qquad Z_i^* = Z_{(i)}/\sqrt{N\theta_i(1+\theta_i)} \qquad (i = 1,2,\dots,m) \quad .$$

<u>Theorem 5.2</u>:(a)For $\theta_m < 1$ the asymptotic $(N \to\infty)$ joint density of the X_i^* is the same joint normal density as in theorem 5.1 with λ_i replaced by θ_i $(i = 1,2,\dots,m)$. In particular, if $\theta_i=\theta < 1 (i=1,2,\dots,m)$, then the asymptotic $(N\to\infty)$ joint density of the X_i^* is

$$(5.17)\quad \frac{1}{\sqrt{1+m\theta}}\left(\frac{1+\theta}{2\pi}\right)^{m/2}\exp\left\{-\frac{(1+\theta)}{2(1+m\theta)}\left[(1+(m-1)\theta)\sum_{i=1}^{m}x_i^2-2\theta\sum_{i<j}x_ix_j\right]\right\}.$$

(b) For $\theta_i = 1 (i = 1,2,\ldots,m)$ the conditional asymptotic $(N\to\infty)$ joint density of the

$$(5.18)\quad Z_i^* = \frac{Z_{(i)}}{\sqrt{2N}} = \frac{N-X_{(i)}}{\sqrt{2N}} = -X_i^* \qquad (i = 1,2,\ldots,m) \quad ,$$

given that cell $C_{(k)}$ empties first, is the restricted joint normal density

$$(5.19)\quad g(z_1,\ldots,z_m) = \sqrt{\frac{k}{\pi^m}}\exp\left\{-\left[m\sum_{i=1}^{m}z_i^2-2\sum_{i<j}z_iz_j\right]\Big/k\right\}$$

where $z_i > 0$ $(i = 1,2,\ldots,m)$; with the appropriate relabeling of the z's, the same result holds if any other cell empties first.

Proof: Setting $x_o' = x_i\sqrt{\delta_1}+\ldots+x_m\sqrt{\delta_m}$ where $\delta_i = \theta_i(1+\theta_i)(i = 1,2,\ldots,m)$ and $\theta_o = \theta_1+\theta_2+\ldots+\theta_m = (1-p_{[k]})/p_{[k]}$ we easily obtain (cf. (4.1))

$$(5.20)\quad P\left\{X_i^* = x_i(i = 1,2,\ldots,m)\right\} = \frac{\Gamma(N+x_o'\sqrt{N}+\theta_oN)}{\Gamma(N)}\quad p_{[k]}^{N}\prod_{i=1}^{m}\left[\frac{p_{[i]}^{x_i\sqrt{N\delta_i}+N\theta_i}}{\Gamma(x_i\sqrt{N\delta_i}+N\theta_i+1)}\right].$$

Since each X_i^* takes steps of size $(N\delta_i)^{-\frac{1}{2}}$, we need to adjoin the factor $(N^m\prod_1^m\delta_i)^{\frac{1}{2}}$ on the right side of (5.20) in passing over to the limiting continuous density. Using Stirling's formula for $\Gamma(\cdot)$ throughout (5.19) and passing to the limit (with the above factor adjoined) gives the first result of theorem 5.2 . The proof of (5.19) is quite similar except that we multiply the final result by k to make it a probability density.

Remark: The vector $(X^*_{(1)}, X^*_{(2)}, \ldots, X^*_{(m)})$ is asymptotically a "bona fide" random vector when $\theta_m < 1$ since $P\{E_{(k)}\} \to 1$ as $N \to \infty$, i.e., $P\{X^*_{(k)} = N\} \to 1$ as $N \to \infty$. We also remark that the Z^*_i in (5.18) are interchangeable, and equicorrelated, and that the corresponding unrestricted chance variables Z'_i(say) with $-\infty < Z'_i < \infty (i = 1,2,\ldots,m)$, whose density is $1/k$ times the density in (5.19), are standardized normal chance variables with common correlation $\rho = \frac{1}{2}$.

Corollary 5.4 : If $\delta_{k,i} = \delta = p/q > 1$, where p and q are defined in (4.4), then the X^*_i in (5.15) are jointly normal with mean 0, variance one and common correlation

$$(5.21) \qquad \rho = \frac{q}{p+q} = \frac{1}{1+\delta} < \frac{1}{2} \quad .$$

Another type of result follows from the fact that the conditional c.d.f. of $X_o = X_{(1)} + X_{(2)} + \ldots + X_{(m)}$, given that cell $C_{(k)}$ empties first, can be written as

$$(5.22) \quad P\left\{X_o \leq x\right\} = \sum_{x_o=0}^{x} \frac{\Gamma(N+x_o)}{\Gamma(N)x_o!} p^N_{[k]} (1-p_{[k]})^{x_o} \sum \frac{x_o!}{\prod_1^m x_i!} \prod_{j=1}^{m} \left[\frac{p_{[j]}}{1-p_{[k]}}\right]^{x_j} ,$$

where the multinomial sum is over m-vectors $(x_1, x_2, \ldots, x_m)$ with $x_1+x_2+\ldots+x_m = x_o$ and $0 \leq x_i \leq N-1 (i=1,2,\ldots,m)$. If we disregard the latter restriction or if $x \leq N-1$ then the multinomial sum in (5.22) is unity and we obtain

$$(5.23) \quad P\left\{X_o \leq x\right\} = p^N_{[k]} \sum_{i=0}^{x} \frac{\Gamma(N+i)}{\Gamma(N)i!} (1-p_{[k]})^i = I_{p_{[k]}}(N, x+1)$$

(cf. equations (2.24) and (2.3) in [7]).

A lower bound to $E_k(X_o) \equiv E(X_o | C_{(k)}$ empties first) is easily seen to be

$$(5.24)\quad E_k(X_o) \geq E_k(X_o | X_o \leq N-1) = \frac{p_{[k]}^N}{I_{p_{[k]}}(N,N)} \sum_{i=0}^{N-1} \frac{i\Gamma(N+i)}{\Gamma(N)i!} (1-p_{[k]})^i$$

$$= \frac{N(1-p_{[k]})}{p_{[k]}} \quad \frac{I_{p_{[k]}}(N+1,N-1)}{I_{p_{[k]}}(N,N)} = \frac{N(1-p_{[k]})}{p_{[k]}} \left[1 - \frac{b_{2N-1}(N;p_{[k]})}{I_{p_{[k]}}(N,N)} \right] ,$$

where $b_n(x;p)$ is the binomial probability $\binom{n}{x}p^x(1-p)^{n-x}$. For $p_{[k]} = 1/k$

$$(5.25)\quad Nm \left[1 - \frac{b_{2N-1}(N;1/k)}{I_{1/k}(N,N)} \right] \leq E_k(X_o | W) \leq m(N-1) \quad ;$$

by symmetry, the same result (5.25) holds whichever cell empties first and hence it also holds unconditionally. Since the ratio of the left and right members of (5.25) tends to one, the two limits in (5.25) provide close bounds for large N.

6. Exact and Asymptotic Evaluation of E(T).

In order to assess the cost of using the ISP and to make comparisons with other procedures, e.g., the FSP, it is necessary to obtain exact and asymptotic expressions for $E(T|R_1)$; the asymptotic being needed because the exact ones are too involved to make direct comparisons.

We define a generalized least favorable (GLF) configuration to be one with

$$(6.1)\quad p_{[1]} = p_{[2]} = \cdots = p_{[m]} = q(\text{say}) \text{ and } p_{[k]} = p(\text{say}) .$$

We are interested in $E(T|GLF)$ and, in particular $E(T|LF)$ and $E(T|W)$, for procedure R_1. First we shall give an upper bound for $E(T|GLF)$ based on a simple probability argument. Then we shall develop exact formulas for $E(T|GLF)$ and asymptotic ($\delta^* \to 1$) approximations based on these exact results.

It will be seen that the leading term in the asymptotic approximations is in agreement with the upper bound.

Upper Bound for $E(T)$

An upper bound to $E(T|GLF)$ for any number of cells k can be obtained by using a negative binomial probability argument. We assume $q < p$ and write

$$(6.2)\quad E(T|GLF) = P^*E(T|GLF), CS) + (1-P^*)\, E(T|GLF, IS)$$

where IS denotes any incorrect selection. Consider the first conditional expectation on the right side of (6.2) and let an obsevation from the cell $C_{(k)}$ be called a success and an observation from any other cell be called a failure. Then the expected number of observations required to obtain the first success is well-known to be $1/p$; hence we obtain N/p for N successes. This argument disregards the restriction we need, viz., that the number of failures from any one cell should be at most N. It is therefore clear that the resulting value will necessarily be an upper bound. A similar argument holds for the second conditional expectation in (6.2) and we thus obtain

$$(6.3)\quad E(T|GLF) \leq P^*\left(\frac{N}{p}\right) + (1-P^*)\left(\frac{N}{q}\right) \quad .$$

Under the GLF-configuration we have $q \leq 1/k$ and hence $N/q \geq kN > kN-m = \text{Max } T$, where the maximum is over all possible configurations. Hence

$$(6.4)\quad E(T|GLF) \leq P^*\left(\frac{N}{p}\right) + (1-P^*)(kN-m),$$

which for the LF configuration becomes

$$(6.5)\quad E(T|LF) \leq P^*N\left(\frac{m+\delta^*}{\delta^*}\right) + (1-P^*)[kN-m] \quad .$$

This upper bound will be used to make comparisons with the FSP in Section 7. The coefficient of P^* in (6.5) can also be regarded as a useful (i.e., quickly computable) rough approximation to $E(T|LF)$ for large N since (for large N and fixed δ^*) P^* will be close to 1 and the restrictions, which were omitted in order to obtain (6.5), have a negligible effect on the result; more precise asymptotic results for $E(T|LF)$ are obtained below.

An inequality on $E(T)$ for the W-configuration can be obtained by adding N to each of the three members of (5.25); this gives

$$(6.6) \qquad N\left[k - \frac{m b_{2N-1}(N,1/k)}{I_{1/k}(N,N)}\right] \leq E(T|W) \leq kN-m \quad .$$

<u>Exact Evaluation of E(T)</u>

Clearly we have for any configuration

$$(6.7) \quad E(T) = N + \sum_{i=1}^{k} P(E_{(i)}) \left[\sum_{j \neq i} E(X_{(j)}|E_{(i)})\right] .$$

For the GLF configuration, using symmetry, we obtain from (6.7)

$$(6.8) \quad E(T|GLF) = N + P(E_{(k)})(k-1)\, E(X_{(1)}|E_{(k)})$$

$$+ (k-1)\left[\frac{1-P(E_{(k)})}{k-1}\right]\left[E(X_{(k)}|E_{(1)}) + (k-2)\, E(X_{(2)}|E_{(1)})\right]$$

$$= N + (k-1)\,[\mu_{1,k} + \mu_{k,1} + (k-2)\,\mu_{2,1}]$$

where for $i \neq j$

$$(6.9) \quad \mu_{i,j} = P(E_{(j)})E(X_{(i)}|E_{(j)}) = \sum_{x=0}^{N-1} x\, P(X_{(i)} = x, E_{(j)})$$

$$= \begin{cases} p^N \sum \dfrac{x_1 \Gamma(N+x_o)}{\Gamma(N) \prod_{\alpha=1}^{m} x_\alpha!} \, q^{x_o} & \text{for } j=k,\ i \neq k \\[2ex] q^N \sum \dfrac{x_1 \Gamma(N+x_o)}{\Gamma(N) \prod_{\alpha=1}^{m} x_\alpha!} \, p^{x_1} q^{x_o - x_1} & \text{for } i=k,\ j \neq k \\[2ex] q^N \sum \dfrac{x_2 \Gamma(N+x_o)}{\Gamma(N) \prod_{\alpha=1}^{m} x_\alpha!} \, p^{x_1} q^{x_o - x_1} & \text{for } i \neq k,\ j \neq k \text{(needed only for } k \geq 3), \end{cases}$$

$x_0=x_1+x_2+\ldots+x_m$ throughout and each summation is over all ordered m-tuples $(x_1,x_2,\ldots,x_m)$ with $0 \leq x_i \leq N-1$ $(i = 1,2,\ldots,m)$. For the LF configuration we merely set $p = \delta^*/(m+\delta^*)$ and $q = 1/(m+\delta^*)$; in the W-configuration we set $p = q = 1/k$ and obtain

$$(6.10)\quad E(T|W) = N + (k-1)\ E(X_{(2)}|E_{(1)}) = N + (k-1)k\ \mu$$

where

$$(6.11)\quad \mu = \sum_{x=0}^{N-1} x\ P(X_{(2)}= x,\ E_{(1)}) = \sum \frac{x_2\Gamma(N+x_0)}{\Gamma(N)\prod_{i=1}^{m} x_i!} \left(\frac{1}{k}\right)^{N+x_0} ,$$

the summation being over the same range as in (6.9). Using the above result (4.2) taken form [7], the above sums in (6.9) and (6.11) can be written in the form of Dirichlet integrals and those that do not have equal exponents and equal lower limits of integration can be "symmetrized" by using appropriate identities and integrations-by-parts. We state the result first in terms of the Dirichlet integral

$$(6.12)\quad D_m(M,N_1,\ldots,N_m;\ a_1,a_2,\ldots,a_m)$$

$$= \frac{\Gamma(M+N_0)}{\Gamma(M)\prod_{x=1}^{m}\Gamma(N_i)} \int_{a_1}^{\infty}\cdots\int_{a_m}^{\infty} \frac{\prod_{i=1}^{m}\left[y_i^{N_i-1}dy_i\right]}{(1+y_0)^{M+N_0}} \qquad \begin{matrix}(M,N_i> 0)\\(i=1,2,\ldots,m)\end{matrix}$$

where $y_0=y_1+y_2+\ldots+y_m$ and $N_0=N_1+N_2+\ldots+N_m$, and finally in terms of "semi-symmetric" Dirichlet integrals with $N_1=\ldots=N_m=N$(say) or "completely symmetric" Dirichlet integrals with $M=N_1=\ldots=N_m=N$(say). We shall write $D_m(M,N_1,\ldots,N_m;a)$ if $a_1=a_2=\ldots a_m=a$ and if, in addition, we have $N_1=N_2=\ldots=N_m=N$ then we write $D_m(M,N;a)$; by definition $D_0(N;a) \equiv 1$ and it is easily checked that $D_1(M,N;a) = I_\theta(M,N)$ where $\theta = (1+a)^{-1}$. For convenience let $q_1=q/(p+q)$.

Theorem 6.1: For the GLF-configuration with $q \le p$

$$(6.13) \quad E(T|GLF) = \frac{N}{q}\left[1-\left(\frac{p-q}{p}\right) D_m(N,N;\, q/p) - \frac{1}{2p}\, b_{2N}(N;q_1)\, D_{m-1}(2N,N;q_1)\right] .$$

[Remark: Note that $q < p$ and $q = p$ are both included in (6.13). Special results for the LF-configuration and for $k=2$ are given at the end of the proof of theorem 6.1. Another expression for $E(T|LF)$ in terms of the derivative of $D_m(N,N;q/p)$ is given in equation (B7) of the Appendix].

For the proof of theorem 6.1 we require the following two lemmas. The first lemma is a generalization of the well-known result $1-I_\theta(M,N) = I_{1-\theta}(N,M)$ for the Incomplete Beta function, which is obtained by setting $m = 1$ in (6.14).

Lemma 6.1: For any $\theta \ge 0$, any $M \ge 1$ and any $N_i \ge 1$ $(i = 1,2,\ldots,m)$

$$(6.14) \quad m\, D_m(M,N_1,\ldots,N_m;\theta,1,\ldots,1) = 1-D_m(N_1,M,N_2,\ldots,N_m;1/\theta).$$

More generally, if θ in the left member of (6.14) is associated with the integral of x_i, which has the exponent N_i-1, then an analogous result holds with N_i as the first argument and M as the $(i+1)^{st}$ argument in the expression corresponding to the right side of (6.14).

Proof: Starting with the right side of (6.14), the region of integration is such that at least one $y_i < 1/\theta$. This integral, by a symmetry argument, is easily seen to be m times the integral with y_m as the smallest y_i, i.e.,

$$(6.15) \quad 1-D_m(M,N_1,\ldots,N_m;1/\theta) = \frac{m\,\Gamma(M+N_o)}{\Gamma(M)\prod_{i=1}^{m}\Gamma(N_i)} \int_0^{\frac{1}{\theta}} \int_{y_m}^{\infty} \cdots \int_{y_m}^{\infty} \frac{\prod_{i=1}^{m}\left[y_i^{N_i-1}dy_i\right]}{(1+y_o)^{M+N_o}} .$$

If we now let $x_i = y_i/y_m$ $(i = 1,2,\ldots,m-1)$ and $x_m = 1/y_m$ then the Jacobian is $(1/x_m)^{m+1}$ and the result (6.14) follows immediately after simplification.

Lemma 6.2: If $N_1=N_2=\ldots=N_m=N$(say) and $M > 1$, then

(6.16) $D_m(M,N;a) = D_m(M-1,N;a) - \frac{Nm}{M+N-1} b_{M+N-1}(N;a_1) D_{m-1}(M+N-1,N;a_1)$

where $a_1 = a/(1+a)$.

Proof: Starting with the left member of (6.16) we insert $(1+y_o) - y_o$ in the integrand and separate into 2 terms. Using symmetry on the 2^d integral, we obtain

(6.17) $D_m(M,N;a) = (\frac{M+m.N-1}{M-1}) D_m(M-1,N;a) - \frac{Nm}{M-1} D_m(M-1,N+1,N,\ldots,N;a)$.

Integrating the last member of (6.17) by parts with respect to y_1 and collecting like terms gives the required result (6.16).

Proof of theorem 6.1: We shall evaluate each of the $\mu_{i,j}$ in the last expression of (6.8). From (4.2) and (6.12)

(6.18) $\mu_{1,k} = \frac{Nq}{p} D_m(N+1,N-1,N,\ldots,N;q/p)$

(6.19) $\mu_{k,1} = \frac{Np}{q} D_m(N+1,N-1,N,\ldots,N;p/q,1,\ldots,1)$

(6.20) $\mu_{2,1} = ND_m(N+1,N-1,N,\ldots,N;1,\ldots,1,p/q)$

Integrating-by-parts with respect to y_1 in (6.18) gives

(6.21) $\mu_{1,k} = \frac{Nq}{p} D_m(N+1,N;q/p) - \frac{N}{2} b_{2N}(N;q_1) D_{m-1}(2N,N;q_1)$.

Using lemma 6.2 to replace D_m above we obtain

(6.22) $\mu_{1,k} = \frac{Nq}{p} D_m(N,N;q/p) - \frac{N}{2p} b_{2N}(N;q_1) D_{m-1}(2N,N;q_1)$.

For (6.19) we first apply lemma 6.1 with $\theta = p/q$, $N_1=N-1, N_2=N_3=\ldots=N_m=N$ and $M=N+1$ and obtain

(6.23) $\mu_{k,1} = \frac{Np}{mq} \; 1-D_m(N-1,N+1,N,\ldots,N;q/p)$

First integrating-by-parts with respect to y_1 and then using lemma 6.2 gives

$$(6.24)\quad \mu_{k,1} = \frac{Np}{mq}\left[1 - D_m(N,N;q/p)\right] - \frac{N(kN-1)}{m(2N-1)}\ b_{2N-1}(N-1;q_1)\ D_{m-1}(2N-1,N;q_1)$$

Applying lemma 6.1 to (6.20) gives

$$(6.25)\quad \mu_{2,1} = \frac{N}{m}\left[1 - D_m(N,N-1,N,\ldots,N,N+1;q/p)\right] \quad .$$

Integrating-by-parts (twice) with respect to y_1 and y_m gives

$$(6.26)\quad \mu_{2,1} = \frac{N}{m}\left[1 - D_m(N,N;q/p) - \tfrac{1}{2}\, b_{2N}(N;q_1)\ D_{m-1}(2N,N;q_1)\right]$$

$$+ \frac{N^2}{m}\ b_{2N-1}(N-1;q_1)\left[\frac{D_m(2N-1,N;q_1)}{2N-1} + b_{3N-1}(N;q_2)\ \frac{D_{m-2}(3N-1,N;q_2)}{3N-1}\right]$$

where $q_2 = q/(1+q)$. If we now use lemma 6.2 with $M = 2N$ and $a_1 = q_2$ to replace the last term in (6.26), then

$$(6.27)\quad \mu_{2,1} = \frac{N}{m}\left[1 - D_m(N,N;q/p)\right]$$

$$+ \frac{N}{m(m-1)}\ b_{2N-1}(N-1,q_1)\left[\frac{kN-1}{2N-1}\ D_{m-1}(2N-1,N;q_1) - \left(\frac{1}{p+q}\right) D_{m-1}(2N,N;q_1)\right]$$

Substituting (6.22), (6.24) and (6.27) in (6.8) gives finally the desired result (6.13); this proves theorem 6.1.

For the special case $k = 2$ we obtain a simpler result for $E(T|GLF)$ with $q \leq p$ in terms of the Incomplete Beta function, i.e., for $N > 1$

$$(6.26)\quad E(T|GLF) = N\left[1 + \frac{q}{p}\ I_p(N+1,N-1) + \frac{p}{q}\ I_q(N+1,N-1)\right]$$

and for $p = q$ we have $p = q = \frac{1}{2}$ and this reduces to

$$(6.27)\quad E(T|W) = N\left[1+2\ I_{\frac{1}{2}}(N+1,N-1)\right] = 2N\left[1-\binom{2N}{N}\left(\frac{1}{2}\right)^{2N}\right].$$

Clearly if $N = 1$, then $T = 1$ and, of course, $E(T) = 1$. It is easily verified that these results are in agreement with (6.13).

Remark: For the more general ISP problem with $N_i \geq 1$ and $p_i > 0$ associated with cell C_i $(i = 1,2,\ldots,k)$ (cf. sections 2, 3), it should be noted that we can still write E(T) in terms of D_m-integrals; in fact, the result by (4.2) is easily seen to be

$$(6.28)\quad E(T) = \sum_{j=1}^{k} N_j\, D_m\left(N_j, N_1,\ldots,N_k; \frac{p_1}{p_j},\ldots,\frac{p_k}{p_j}\right)$$

$$+ \sum_{j=1}^{k} \sum_{\substack{i=1\\ i\neq j}}^{k} \frac{N_j p_i}{p_j}\, D_m\left(N_j+1, N_1,\ldots,N_i-1,\ldots,N_k; \frac{p_1}{p_j},\ldots,\frac{p_k}{p_j}\right),$$

where it is to be understood that, for each j, N_j is only in the first argument of D_m and the lower limit of integration p_j/p_j is omitted. This can be further simplified using a generalization of lemma 6.2, but lemma 6.1 is not available for unequal p's and the resulting D_m's are not simple; we omit these results as they are not used here.

Corollary 6.1: In addition to the exact $E(T|W)$ obtained by setting all the p's equal and to the exact $E(T|LF)$ obtained by using (4.4), we also have the approximation ($\simeq$)

$$(6.29)\quad E(T|LF) \simeq \frac{N}{\delta^*}(m+\delta^*)^2\left\{\frac{[P^*+\delta^*(1-P^*)]}{m+\delta^*} - \frac{1}{2}\, b_{2N}(N;q_1^*)\, D_{m-1}(2N,N;q_1^*)\right\},$$

where $q_1^* = 1/(1+\delta^*)$.

Proof: By the defintion of N in (4.5) we can replace the first D_m-expression by P^* and (6.29) readily follows; the approximation is due only to the discreteness of N.

The following recursion formula is a corollary of lemmas 6.1 and 6.2; it is useful for computing $D_m(M,N;a)$ for small values of m, especially when M is not too large.

Corollary 6.2; For $M \geq 1$, $N \geq 1$, and $m \geq 1$

$$(6.30)\quad D_m(M,N;a) = 1-Nm \sum_{j=0}^{M-1} \frac{b_{N+j}(N;a_1)}{N+j} D_{m-1}(N+j,N;a_1)$$

$$= 1-m\, a_1(1-a_1) \sum_{j=0}^{M-1} b_{N-1+j}(N-1,a_1) D_{m-1}(N+j,N;a_1) \quad .$$

In particular, for m = 2 this gives

$$(6.31)\quad D_2(M,N;a) = 1 - 2\, a_1(1-a_1) \sum_{j=0}^{M-1} b_{N-1+j}(N-1;a_1)\, I_{\overline{a}_2}(N+j,N)$$

where $a_1 = a/(1+a)$ and $\overline{a}_2 = 1/(1+a_1)$.

Proof: Iterating lemma 6.1 on the argument M gives

$$(6.32)\quad D_m(M,N;a) = D_m(1,N;a) - Nm \sum_{j=1}^{M-1} \frac{b_{M+N-j}(N;a_1)}{M+N-j} D_{m-1}(M+N-j,N;a_1) \quad .$$

From lemma 6.2 with $N_1 = 1$ and $M = N_2 = \ldots = N_m = N$, we obtain after integration

$$(6.33)\quad D_m(1,N;a) = 1 - m\, a_1^N D_{m-1}(N,N;a_1)$$

and substituting this in (6.32) gives the desired result (6.30).

Asymptotic Results for E(T).

The results of corollary 5.2, Stirling's approximation to the Γ-function and some lemmas below will be combined to give an asymptotic expression for $E(T|GLF)$ which requires that N be large and that $p > q$. In a subsequent corollary we replace N in this expression as a function of δ^*(for fixed P^*) using (5.14) and, staying in the LF-configuration, we then let $\delta^* \to 1$ (which implies that $N \to \infty$) and obtain an asymptotic approximation for $E(T|LF)$. Later we also obtain asymptotic ($N \to \infty$) results for $E(T|W)$.

Theorem 6.2: For the GLF-configuration with $q < p$ and N large

$$(6.34) \quad E(T|GLF) \sim \frac{N}{q}\left\{1 - \left(\frac{p-q}{p}\right)\int_{-\infty}^{\infty} F^m\left(x + \sqrt{N}\log\frac{p}{q}\right) f(x)dx \right.$$

$$\left. - \frac{1}{2p\sqrt{N\pi}} \left[\frac{4pq}{(p+q)^2}\right]^N \int_{-\infty}^{\infty} F^{m-1}\left(\frac{x+\sqrt{N}\log\left(\frac{p+q}{2q}\right)}{\sqrt{2}}\right) f(x)dx\right\}$$

Proof: By corollary 5.2 with $\lambda_i = \lambda = \lim(N/M)$ as $N\to\infty (i = 1,2,\ldots,m)$ and by an argument similar to the one used for corollary 5.1, we obtain for $q < p$

$$(6.35) \quad D_m(M,N;a) \sim \int_{-\infty}^{\infty} F^m\left(x\sqrt{\lambda} + \sqrt{N\lambda}\log\frac{\lambda}{a}\right) f(x)dx \; .$$

Applying this in (6.13) for D_m and D_{m-1} and using Stirling's approximation to $\Gamma(\cdot)$ in the binomial factor, gives the desired result (6.34).

Corollary 6.2: For the LF-configuration with N large

$$(6.36) \quad E(T|LF) \sim N\left(\frac{m+\delta^*}{\delta^*}\right) \left[P^* + \delta^*(1-P^*)\right]$$

$$- \sqrt{\frac{N}{\pi}}\,\frac{(m+\delta^*)^2}{2\delta^*}\left[\frac{4\delta^*}{(1+\delta^*)^2}\right]^N \int_{-\infty}^{\infty} F^{m-1}\left(\frac{x+\sqrt{N}\log\left(\frac{1+\delta^*}{2}\right)}{\sqrt{2}}\right) f(x)dx \; .$$

Proof: This follows from (6.34), (5.7) and (5.13) with $\lambda = 1$.

We shall use the symbol $E_o(T|LF)$ to denote the limit of $E(T|LF)$ as $\delta^*\to 1$, after using (5.14) to replace N; let $H = H_m(P^*,\frac{1}{2})$ be defined by (5.7).

As $\delta^*\to 1$ (for fixed P^*) in the LF-configuration, it follows from (5.14) that $N\to\infty$ and expanding (6.36) in increasing powers of $\delta^* - 1$ gives

(6.37) $$E_o(T|LF) \sim \frac{2H^2}{(\log \delta^*)^2} \; [k + (\delta^*-1)(1-kP^*)]$$

$$- \frac{k^2H\ f(H)}{\log \delta^*} \int_{-\infty}^{\infty} F^{m-1}\left(\frac{x\sqrt{2}+H}{2}\right) f(x)dx + \mathcal{O}\left(\frac{1}{\log \delta^*}\right)$$

$$\sim \frac{2kH^2}{(\delta^*-1)^2} + \frac{H}{\delta^*-1}\left\{ 2H[1+k(1-P^*)]\right.$$

$$\left. - k^2f(H)\int_{-\infty}^{\infty} F^{m-1}\left(\frac{x\sqrt{2}+H}{2}\right) f(x)dx\right\} + \mathcal{O}\left(\frac{1}{\delta^*-1}\right) .$$

Thus for δ^* sufficiently close to 1 we have

(6.38) $$E_o(T|LF) \leq \frac{2kH^2}{(\delta^*-1)^2} + \frac{2H^2}{\delta^*-1} \; [1 + k(1-P^*)] .$$

It is interesting to compare (6.38) with the corresponding result obtained from the upper bound (6.5); expanding the latter about $\delta^* = 1$ gives

(6.39) $$E_o(T|LF) \leq \frac{2kH^2}{(\delta^*-1)^2} + \frac{2H^2}{\delta^*-1} \; [P^* + k(1-P^*)] ,$$

which gives a better upper bound than (6.38). We shall therefore use (6.39) in making comparisons with the FSP in Section 7.

For the W-configuration (i.e., with equal parameters) we use (6.13) with $q = p$ and $N\to\infty$ to obtain an asymptotic expression for $E(T|W)$. This is the direct approach involving the limit (of the exact expression for $E(T|W)$) as $N\to\infty$ and is somewhat different from that obtained by finding $E(T|W)$ for the limiting distribution, i.e., for the "resticted normal" distribution. The former, i.e., the limit of the exact $E(T|W)$, is the more appropriate value for our needs and it is also numerically more meaningful; the latter is given without proof for purposes of comparison. More specifically, these two asymptotic results agree on the "leading term", kN, but give different results for the "$\sqrt{N}$ term"; this points up the fact that in general the limit of the exact expectation need not be the same as the expectation under the limiting distribution.

Corollary 6.4: For the W-configuration

$$(6.40)\quad E(T|W) \sim kN \left[1 - \frac{k}{2\sqrt{N\pi}} \int_{-\infty}^{\infty} F^{m-1}\left(\frac{x}{\sqrt{2}}\right) f(x)dx\right]$$

Proof: Clearly we can disregard the middle term in (6.13) and using Stirling's approximation for $\Gamma(\cdot)$ we obtain $(N\pi)^{-\frac{1}{2}}$ for the binomial factor. For the last factor, we consider upper and lower bounds by writing

$$(6.41)\quad D_{m-1}(2N,N;q_1e^{\epsilon/N}) \le D_{m-1}(2N,N;q_1) \le D_{m-1}(2N,N;q_1e^{-\epsilon/N})$$

for a fixed $\epsilon > 0$. Since $D_{m-1}(2N,N;q_1)$ must lie between these for every N, it is easy to see using corollary 5.2 and continuity with respect to ϵ in (6.41) that $D_{m-1}(2N,N;q_1)$ approaches the orthant probability

$$(6.42)\quad D_{m-1}(2N,N;q_1) \sim \int_{-\infty}^{\infty} F^{m-1}\left(\frac{x}{\sqrt{2}}\right) f(x)dx \,,$$

which completes the proof of (6.40). The computations of numerical values for $E(T|W)$ in Table II are based on (6.40); the values of the right side of (6.42) are tabulated in [8] and the notation used there for it is $\overline{V}_{m-1,m-1}(3) = \overline{u}_{m-1}(3)$.

For purposes of comparison we now give (without proof) formulas for $E(T|W)$ and $\sigma^2(T|W)$ computed under the limiting "restricted normal" distribution (cf. theorem 5.2): they are

$$(6.43)\quad E(T|W) \sim k[N - \sqrt{N}\, E(V_k)]$$

$$(6.44)\quad \sigma^2(T|W) \sim N[(m+1)^2\, \sigma^2(V_k) - 2m] \,,$$

where V_k is the maximum of k independent standard normal chance variables.

7. Comparison of ISP and FSP

In this section the ISP will be conpared with the FSP which selects the cell with the highest observed frequency as the best cell and uses randomization when ties occur. A separate comparison will be made for $k = 2$ and for $k > 2$, since the results for $k = 2$ form one of the motivating factors for writing this paper. One of the difficulties of making "small sample" comparisons of the ISP and FSP is that the tables in [1] give the values of the P(CS) for fixed $n_o = 1(1)30$ instead of giving the values of n for fixed P*; this makes them quite inadequate for any triple (k, δ^*, P^*) that requires $n > 30$. There is a large sample approximation given but no table gives explicit n_o values based on the normal approximation. Special values were extracted from the tables in [1] and numerical comparisons are made in Table II for these selected values. Asymptotic $(\delta^* \to 1)$ comparisons for $k \geq 2$ given below indicate that the ISP is more efficient than the FSP for any $P^* \geq \frac{1}{2}$; for $k = 2$ we can state a stronger result, i.e., that the random variable T is at most equal to n_o with probability 1.

Binomial Case (k = 2)

The $P(CS|R_1)$ for the ISP with $k = 2$ is given in (4.6). For the FSP, which we denote by R_o, the corresponding $P(CS|R_o)$ for $k = 2$ can also easily be written as an Incomplete Beta function and we obtain for odd n

$$(7.1)\quad P(CS|R_o) = \sum_{i=\frac{n+1}{2}}^{n} \binom{n}{i} p^i q^{n-i} = I_p\left(\frac{n+1}{2}, \frac{n+1}{2}\right) ,$$

and for even n

$$(7.2)\quad P(CS|R_o) = \tfrac{1}{2}\binom{n}{n/2} p^{n/2} q^{n/2} + \sum_{i=\frac{n}{2}+1}^{n} \binom{n}{i} p^i q^{n-i}$$

$$= \tfrac{1}{2}\left[I_p\left(\frac{n}{2}, \frac{n}{2}+1\right) + I_p\left(\frac{n}{2}+1, \frac{n}{2}\right)\right] = I_p\left(\frac{n}{2}, \frac{n}{2}\right) ,$$

where $p \geq \frac{1}{2}$ and the last equality of (7.2) is straightforward. It follows from (7.1) and (7.2) that we get the same probability for any odd n and the even integer n + 1, and this is evident in the tables for k = 2 in [1]; hence we can restrict our attention to odd integer solutions.

It follows that the value of n required by the FSP for k = 2 is the smallest odd integer $n_o \geq x$ where x is the solution of

$$(7.3) \quad I_{\delta^*/(1+\delta^*)}\left(\frac{x+1}{2}, \frac{x+1}{2}\right) = P^* \quad .$$

Comparing (7.7) with the corresponding probability requirement based on (4.6) with the same δ^* and P^*, we get the astonishing result that

$$(7.4) \quad n_o = 2N-1 = \text{Max } (T) > E(T) \quad ,$$

where the last inequality holds for any configuration of the parameters.

We define a procedure R to be more efficient relative to (δ^*, P^*) than R' if they satisfy the same probability requirement and the $E(T|R) \leq E(T|R')$ in the LF-configuration; if the inequality holds for all pairs (δ^*, P^*) then we say that R is uniformly more efficient than R'.

Since (7.4) holds for any configuration, it follows a fortiori that we have proved

Theorem 7.1: For k = 2 with any δ^* and any P^* the ISP is uniformly more efficient than the FSP.

Moreover for k = 2 we can assert the stronger result that with probability one the random variable T associated with the ISP is at most equal to the value of n_o required by the FSP for the same probability requirement; it should be pointed out that this stronger result does not hold for k > 2.

Multinomial Case $(k > 2)$

As already mentioned exact "small sample" comparisons are difficult to make for $k > 2$ and it is necessary to make "asymptotic" comparisons. Our principal results deal with the LF- and W-configurations with P* fixed and $\delta^* \to 1$. In accord with the definition of "more efficient" above, we now define for fixed P* the first order asymptotic proportion saved for any procedure R relative to R' as

$$(7.5) \quad APS_1(R,R'; P^*) = \lim_{\delta^* \to 1} \left[\frac{E(T|R',LF)}{E(T|R,LF)} - 1\right] .$$

Of course, $E(T|LF) \equiv n_o$ for the FSP; we shall take R_o for the procedure R' and R_1 for R in (7.5). If the limit in (7.5) is zero then we define the corresponding second order asymptotic proportion saved for R relative to R' as

$$(7.6) \quad APS_2(R,R'; P^*) = \lim_{\delta^* \to 1} \frac{1}{(\delta^*-1)} \left[\frac{E(T|R',LF)}{E(T|R,LF)} - 1\right] .$$

Theorem 7.2: For $k \geq 2$ and all P*

$$(7.7) \quad APS_1(R_1,R_o; P^*) = 0 ,$$

$$(7.8) \quad APS_2(R_1,R_o; P^*) \geq \frac{k+2}{2[k-(k-1)P^*]} - 1 ,$$

which is positive for all $P^* \geq \frac{1}{2}$; in particular, for k=2 and 3, it is positive for all non-trivial P*-values, i.e., for all $P^* > 1/k$.

Proof: It is clear that for fixed P* as $\delta^* \to 1$, $N \to \infty$. For R_1 we shall use the results of (6.37) and (6.39), but first we need analogous results for the $n_o \equiv E(T|R_o)$ of the FSP.

Using the arc sine-square root transformation and the subsequent normal approximation for the FSP given in [1], we have asymptotically $(\delta^* \to 1)$, using the notation in [1],

$$(7.9)\quad n_o \sim \frac{b_1}{a_1^2}\ \frac{\Lambda^2}{2} = \frac{b_1}{a_1^2}\ H_m^2\,(P^*,\tfrac{1}{2}) \quad ,$$

where we have set $\rho = \frac{1}{2}$ since $\rho \to \frac{1}{2}$ as $\delta^* \to 1$, and where $a_1 = 2\left[\arc \sin\sqrt{p} - \arc \sin\sqrt{q}\right]$,

$$(7.10)\quad b_1 = 2 + 2\sqrt{\frac{p}{m(1-q)}} \quad ,$$

and p, q are in the LF-configuration (4.4). Expanding the right side of (7.9) in increasing powers of $\delta^* - 1$ we obtain

$$(7.11)\quad n_o \sim \frac{2kH_m^2(P^*,\frac{1}{2})}{(\delta^* - 1)^2} + \frac{(k+2)\ H_m^2(P^*,\frac{1}{2})}{\delta^* - 1} + \mathcal{O}(1) \quad .$$

Using the first terms of (6.37) and (7.11) we obtain (7.7). To get a lower bound for $APS_2(R_1,R_o;P^*)$ we use the upper bound given by (6.39) and (7.11) to obtain the right member of (7.8); the latter is positive whenever

$$(7.12)\quad P^* > \tfrac{1}{2}\left(\frac{k-2}{k-1}\right)$$

and, a fortiori, whenever $P^* > \frac{1}{2}$. Since the right side of (7.12) is $\leq 1/k$ for k = 2 and 3, the rest of the assertion in theorem 7.2 follows.

It is also of interest to see whether the stronger result, that $E(T|W) \leq n_o$, is true for any values of $k > 2$; we have already seen that this holds for k = 2. For this purpose we expand (6.40) in increasing powers of $\delta^* - 1$ after substituting for N from (5.14) and compare term-by-term with the expansion in (7.11); the values of the integral in (6.38) are taken from [8]. We find that for δ^* sufficiently close to 1 (and hence N large)

$$(7.13)\quad \begin{aligned} E(T|W) &< n_o \qquad \text{for } k \leq 3 \quad , \\ E(T|W) &> n_o \qquad \text{for } k > 3 \quad . \end{aligned}$$

The numerical results of Table II are consistent with (7.13).

APPENDIX

A: Remarks on the Variance of T.

It is interesting to note that by applying (4.2) we can obtain an exact expression for ET^2, and hence for $\sigma^2(T)$, in terms of D_m- integrals defined in (6.12). This initial expression can then be symmetrized by using lemmas 6.1 and 6.2 but the final expression appears to be "tediously" long and is omitted. The intermediate result with a common lower limit of integration q/p for each D_m is

(A1)
$$\begin{aligned} E(T^2|GLF) = (2N+1)E(T|GLF) + N(N+1)\Big\{ -1 + \frac{mq^2}{p^2}\Big[D_m(N+2,N-2,N,\dots,N;q/p) \\ + (m-1)D_m(N+2,N-1,N-1,N,\dots,N;q/p)\Big] + (\tfrac{p}{q})^2\Big[1-D_m(N-2,N+2,N,\dots,N;q/p)\Big] \\ + (m-1)\Big[1-D_m(N,N-2,N+2,N,\dots,N;q/p)\Big] + 2(m-1)\frac{p}{q}\Big[1-D_m(N-1,N+2,N-1,N,\dots,N;q/p)\Big] \\ + (m-1)(m-2)\Big[1-D_m(N,N-1,N-1,N+2,N,\dots,N;q/p)\Big]\Big\}. \end{aligned}$$

Using the symmetrized form of (A1) and the result for $E(T|GLF)$ in (6.13) to compute $\sigma^2(T)$, it was found that the sum of the terms, say $T_1(N)$, involving N^2 (after much calculation!) reduced to

(A2)
$$T_1(N) = N^2(p-q)^2D_m(N,N;q/p)\Big[1-D_m(N,N;q/p)\Big]/p^2q^2.$$

It should be noted that $D_m(N,N;q/p) \to 1$ as $N \to \infty$ (or $P^* \to 1$) and $N^2[1-D_m(N,N;q/p)] \to 0$, since $1-D_m(N,N;q/p) \sim 1-P^*$ and it can be shown using (5.14) and the definition of $H_m(P^*, \frac{1}{2})$ in (5.7) that $N \sim -4 \log (1-P^*)$. It is conjectured that a similar result holds for the terms involving $N^{3/2}$ and that the leading "non-zero" term in the asymptotic expression for $\sigma^2(T)$ is the same as that obtained from theorem 5.2(a), namely, for $q < p$

(A3)
$$\sigma^2(T) \sim Nmq/p^2 \quad .$$

It should also be noted that $T_1(N) \equiv 0$ for $q = p$. It is conjectured that a similar result holds for the terms involving $N^{3/2}$ and that the leading "non-zero" term for $\sigma^2(T)$ when $q = p$ is asymptotically equivalent to that given in (6.44).

B: A Property of $D_m(M,N;\theta)$.

For convenience, we denote $D_m(M,N;\theta)$ by $f(M,\theta)$ and we use $M^{(r)}$ to denote the product of the r factors $M(M+1) \ldots (M+r-1)$. Differentiating the expression (6.12) for $D_m(M,N;\theta)$ with respect to θ and letting $D = \partial/\partial\theta$ denote the partial derivative with respect to θ, we obtain

(B1) $$\theta D f(M,\theta) = M\left[f(M+1,\theta) - f(M,\theta)\right] = M \Delta f(M,\theta) \quad ,$$

where Δ denotes the usual finite-difference operator.

We now prove the more general

Theorem A1: For any M, N, θ and any r $(r = 0,1,\ldots)$

(B2) $$M^{(r)}\Delta^r f(M,\theta) = \theta^r D^r f(M,\theta) \quad .$$

Proof. The proof is by induction and we have already shown it to be true for $r = 1$. Assuming it is true for $r = k$, we have by applying the operator $M\Delta$ to the left side of (B2) for $r = k$

(B3) $$\begin{aligned} M\Delta\left\{M^{(k)}\Delta^k f(M,\theta)\right\} &= M\left[(M+1)^{(k)}\Delta^{k+1} f(M,\theta) + k(M+1)^{(k-1)}\Delta^k f(M,\theta)\right] \\ &= M^{(k+1)}\Delta^{k+1} f(M,\theta) + kM^{(k)}\Delta^k f(M,\theta) \end{aligned}$$

Now, using the induction hypothesis and the fact that Δ and D operate on different arguments and hence commute, we have from (B3)

(B4) $$\begin{aligned} M^{(k+1)}\Delta^{k+1} f(M,\theta) + k\theta^k D^k f(M,\theta) &= M\Delta\{\theta^k D^k f(M,\theta)\} \\ &= \theta^k D^k\{M\Delta f(M,\theta)\} = \theta^k D^k\{\theta D f(M,\theta)\} \\ &= \theta^k\left[\theta D^{k+1} f(M,\theta) + kD^k f(M,\theta)\right] \\ &= \theta^{k+1} D^{k+1} f(M,\theta) + k\theta^k D^k f(M,\theta) \end{aligned}$$

from which (B2) follows.

As a corollary we note that if we let $f^{(r)}(M,\theta)$ denote the left (or right) side of (B2) then

(B5) $$M\Delta f^{(r)}(M,\theta) = \theta D f^{(r)}(M,\theta) \qquad (r = 0, 1, \ldots) .$$

We remark that by using lemma (6.2) or by taking the derivative of $D_m(M,N,\theta)$ directly that

(B6) $$\frac{\partial}{\partial\theta} D_m(M,N,\theta) = - \frac{mNM}{\theta(M+N)} b_{M+N}\left(N;\frac{\theta}{1+\theta}\right) D_{m-1}\left(M+N,N,\frac{\theta}{1+\theta}\right) ,$$

and hence the expression for $E(T|GLF)$ in theorem 6.1 can also be written as

(B7) $$E(T|GLF) = \frac{N}{q}\left[1 - \frac{p-q}{p} D_m(N,N;q/p)\right] + \frac{1}{mp^2}\frac{\partial}{\partial\theta} D_m(N,N;q/p) .$$

Acknowledgement

The authors wish to thank Mr. Satindar Kumar, Mr. Desu Mahamanulu and Mr. George G. Woodworth for their help with the computation of Table II.

Table I

N-Values needed to satisfy (4.5) for the ISP.#

P* = .75

δ* \ m	1	2	3	4	5	6	7	8	9
1.2	29^{-1}	63	87	104	119	130	140	149	157
1.4	9^{-0}	19^{-0}	27^{-0}	32	36	40	43	46	48
1.6	5	10	14	17	19	21	23	24^{-0}	26^{-0}
1.8	3	7	10	11	13	14	15	16	17
2.0	3	5	7	9	10	11	11	12	13
2.2	2	4	6	7	8	8	9	10	10
2.4	2	4	5	6	7	7	8	8	8
2.6	2	3	4	5	6	6	7	7	7
2.8	2	3	4	4	5	5	6	6	6
3.0	1	3	3	4	5	5	5	6	6

P* = .90

δ* \ m	1	2	3	4	5	6	7	8	9
1.2	99	151	182	205	223	238	250	261	270
1.4	30	45	55	62	67	71	75	78	81
1.6	16	24^{-0}	29	32	35	38	40	41	43
1.8	10	16	19	21	23	25	26	27	28^{-0}
2.0	8	12	14	16	17	18	19	20	21
2.2	6	9	11	12	13	14	15	16	16
2.4	5	8	9	10	11	12	12	13	13
2.6	4	7	8	9	10	10	11	11	12
2.8	4	6	7	8	8	9	9	10	10
3.0	4	5	6	7	8	8	8	9	9

See footnote on next page.

Table I' (cont.)

P* = .95

δ* \ m	1	2	3	4	5	6	7	8	9
1.2	163	222	257	283	302	318	332	344	354
1.4	49	66	77	84	90	95	99	103	106
1.6	25	34	40	44	47	50	52	54	56
1.8	17	22^{-0}	26^{-0}	29	31	33	34	35	37
2.0	12	16	19^{-0}	21^{-0}	23^{-0}	24^{-0}	25^{-0}	26^{-0}	26^{+1}
2.2	10	13	15	17	18	19	19	20	21
2.4	8	11	12	14	15	15	16	17	17
2.6	7	9	11	12	12	13	14	14	15
2.8	6	8	9	10	11	11	12	12	13
3.0	5	7	8	9	10	10	11	11	11

P* = .99

δ* \ m	1	2	3	4	5	6	7	8	9
1.2	326	395	435	464	486	505	520	533	545
1.4	96	117	129	138	144	150	154	158	162
1.6	51	60	67	71	75	78	80	82	84
1.8	33	39	43	46	49	51	52	54	55
2.0	24	28^{-0}	32	34	36	37	38	39	40
2.2	19	22^{-0}	25^{-0}	26^{+1}	28	29	30	31	32
2.4	15	18^{-0}	20^{-0}	22^{-0}	23^{-0}	24^{-0}	24^{+1}	25^{+1}	25^{+1}
2.6	13	16	17	18	19	20	21	21	22
2.8	11	14	15	16	17	17	18	18	19
3.0	10	12	13	14	15	15	16	16	17

#All entries above the heavy line are based on the asymptotic normal approximation (5.14); those below the heavy line are exact values taken from Table 3 of [4]. The "exponents" in certain boundary cells are the differences between the asymptotic (A) and the exact (E) values of N, i.e., A-E, and give some indication of their agreement.

Table II

Exact and Asymptotic E(T)-Values for the ISP and Comparisons with n_o-Values for the FSP.

δ^*	N, E(T) and n_o-Values		P* = .75, m = 1	P* = .75, m = 2	P* = .75, m = 3	P* = .90, m = 1	P* = .90, m = 2	P* = .90, m = 3
1.2	N		29	(63)	(87)	(99)	(151)	(182)
	E(T|LF)	Exact	50.97	168.23	---	180.15	403.59	---
		Asympt.	(51.01)	(169.89)	(309.84)	(180.15)	(405.29)	(642.04)
	E(T|W)	Exact	51.95	179.15	---	186.77	437.62	---
		Asympt.	(51.92)	(178.92)	(335.20)	(186.77)	(437.45)	(709.48)
	n_o		(57)	(180)	(327)	(198)	(437)	(694)
1.6	N		5	10	14	16	24	29
	E(T|LF)	Exact	7.30	22.86	45.88	25.51	54.71	86.54
		Asympt.	(7.32)	(24.45)	(42.71)	(25.57)	(56.73)	(85.67)
	E(T|W)	Exact	7.54	26.17	47.17	27.52	65.95	102.64
		Asympt.	(7.48)	(25.99)	(50.86)	(27.49)	(65.78)	(109.08)
	n_o		11	26	(46)	31	(64)	(98)
2.0	N		3	5	7	8	12	14
	E(T|LF)	Exact	3.96	10.40	18.93	11.68	25.65	36.40
		Asympt.	(4.08)	(11.02)	(19.63)	(11.80)	(25.76)	(36.78)
	E(T|W)	Exact	4.13	12.37	21.87	12.86	31.78	47.17
		Asympt.	(4.05)	(12.16)	(24.37)	(12.81)	(31.60)	(50.86)
	n_o		5	12	20	15	28	(56)
3.0	N		1	3	3	4	5	6
	E(T|LF)	Exact	1	6.19	7.28	5.16	8.94	13.20
		Asympt.	(0.87)	(6.20)	(7.50)	(5.45)	(9.92)	(13.33)
	E(T|W)	Exact	1	7.02	8.04	5.81	12.37	17.72
		Asympt.	(0.87)	(6.80)	(9.62)	(5.74)	(12.16)	(20.64)
	n_o		1	5	8	7	11	16

All entries in parentheses are asymptotic; all entries without parentheses are exact, except that E(T|LF) for m = 2,3 is based on the approximation (6.29). A "dashed" entry indicates a value not computed. The 6 entries in each cell are based on (4.5) or (5.14), (6.13) or (6.29), (6.36), (6.13), (6.40), and [1] or (7.9), respectively. The second decimal of the E(T)-entries above may not be accurate in some cases.

REFERENCES

[1] Bechhofer, R., Elmaghraby, S. and Morse, N. (1959). A single-sample multiple-decision procedure for selecting the multinomial event which has the highest probability. Ann. Math. Statist. 30 102-119.

[2] Cramer, H. (1946). Methods of Mathematical Statistics. Princeton University Press.

[3] Feller, W. (1957). An Introduction to Probability Theory and its Applications. John Wiley, New York.

[4] Gupta, S. and Sobel, M. (1962). On the smallest of several correlated F-statistics. Biometrika 49 509-523.

[5] Kesten, H. and Morse, N. (1959). A property of the multinomial distribution. Ann. Math. Statist. 30 120-127.

[6] Milton, R. (1963). Tables of the equally correlated multivariate normal probability integral. Technical Report No. 27, Department of Statistics, University of Minnesota.

[7] Olkin, I. and Sobel, M. (1965). Integral expressions for tail probabilities of the multinomial and negative multinomial distributions. Biometrika (to appear in June, 1965).

[8] Ruben, H. (1954). On the moments of order statistics in normal samples. Biometrika 41 200-227.